

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:43:15 ; Search time 210.42 Seconds

(Without alignments)
108,227 Million cell updates/sec

Title: US-09-331-631A-1

Perfect score: 3542

Sequence: 1 MAINTSNLCSDLFLFLSLFL.....SPRSTKQOQPLVSIIDFVGF 666

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/AA1960.DAT.*
2: /SIDSI/gcgdata/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/AA2000.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3542	100.0	666	19	Macadamia integrifolia
2	3412	96.3	666	19	Macadamia integrifolia
3	3227	91.1	625	13	Macadamia integrifolia
4	1137	31.8	566	13	Macadamia integrifolia
5	1103.5	31.2	590	19	Gossypium hirsutum
6	1025	28.9	525	19	Theobroma cacao an
7	898	25.4	593	19	Zea mays antimicro
8	864.5	24.4	637	19	Hordeum vulgare an
9	849.5	24.0	626	20	Peanut allergen A
10	849.5	24.0	626	20	Peanut allergen A
11	849	24.0	614	18	Peanut allergen Ar
12	849	24.0	614	19	Arachis hypogaea a

13	843.5	23.8	626	18	W22150
14	840.5	23.7	605	19	W62838
15	831.5	23.5	605	20	Y40999
16	808	22.8	524	20	W90339
17	806.5	22.8	489	20	W90341
18	756	21.3	444	20	W90340
19	707.5	20.0	415	20	Y40913
20	700.5	19.8	409	20	W90342
21	512.5	14.5	371	20	Y40914
22	238	6.7	611	20	Y29039
23	216.5	6.1	1898	20	Y30795
24	211.5	6.0	1162	21	Y58500
25	195.5	5.5	562	16	R70491
26	195.5	5.5	1239	20	Y55931
27	191	5.4	412	17	W03626
28	190	5.4	360	17	W03627
29	180	5.1	472	15	R47127
30	180	5.1	499	9	P82755
31	171	4.8	1299	21	Y58633
32	168	4.7	510	20	Y15246
33	167	4.7	512	20	Y40912
34	166	4.7	493	13	R26944
35	164.5	4.6	1135	21	Y68784
36	161.5	4.6	1233	20	Y55954
37	158.5	4.5	316	13	R26941
38	158	4.5	1326	20	Y55933
39	157.5	4.4	515	15	R47128
40	156.5	4.4	395	17	W03474
41	155.5	4.4	968	20	Y55966
42	155.5	4.4	1382	18	W31867
43	153.5	4.3	1041	20	W30613
44	148.5	4.2	911	20	Y55957
45	148.5	4.2	1343	18	W31866

ALIGNMENTS

RESULT 1	
ID W62828	W62828 standard; Protein: 666 AA.
AC W62828;	
XX	
DT 27-OCT-1998 (first entry)	
XX	
DE Macadamia integrifolia antimicrobial protein.	
XX	
KW antimicrobial protein; Infestation; control.	
XX	
OS Macadamia integrifolia.	
XX	
FH	
FT key	Location/Qualifiers
FT Peptide	1..28
FT	/note= "signal peptide"
FT Protein	29..666
FT	/note= "mature protein"
XX	
PN W09827805-A1.	
XX	
PD 02-JUL-1998.	
XX	
PF 22-DEC-1997; 97WO-AU00874.	
XX	
PK 20-DEC-1996; 96AU-0004275.	
XX	
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
XX	
PI Bower NT, Goulter KC, Green JL, Manners JM, Marcus JP.	
DR WPI: 1998-377279/32.	
XX	
DR N-PSDB: V42310.	
XX	

Peanut allergen Ar
Glycine max antimi
Soybean beta-congl
G. max Sbp1 protei
G. max Sbp2 protei
G. max truncated S
Ara h 1 allergen P
G. max truncated S
Phaseolin A chain
T. gondii immunoge
A human trichothal
HNB8 ORF 73 protei
Leucocytozoan prot
Human ZC1 protein.
Human thyrotropin
Human follicle sti
Modified oat globi
Rice storage prote
Protein regulating
Peanut allergen, A
Ara h 3 allergen s
P.falciptarum LSA g
Amino acid sequenc
Mouse STE20-relate
P.falciptarum LSA-R
Human ZC3 protein.
Oat globin A2B. A
Mouse SRV-related
Full length human
Human metastasis-a
Human CDC28-#3 RNA
Mouse STE20-relate
Mouse metastasis-a

PT Novel anti-microbial protein from e.g. *Macadamia integrifolia* -
PT useful for controlling microbial infestations of plants or mammals
XX
XX
PS Claim 1; Page 34-36; 96pp; English.
XX

CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
XX
XX

Sequence 666 AA:

Query Match 100.0%; Score 3542; DB 19; Length 666;
Best Local Similarity 100.0%; Pred. No. 2e-299;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAINTSNLCSLLFLSLFLSTTVSLAESEFPDQREYEBCRQCMOLETSGQMRRCVSQCD 60
DB 1 maintsnlcsllflslflstvtslaeseftdqyeecrkqcmqltsgqmrtrcvsgcd 60
OY 61 KREEDIDMSKYDNOEDPQTECCQCRRCROESGPRQOQYCCORCKEICEEEYENRQR 120
DB 61 krtfeedidmskydngedpqtccqqrtrcrqesgprgqyqqrckelceeeeynqr 120
OY 121 DPQOQTECCQKHCQRRETPRHMQTCQRCERRYEKKRQOKRYEQOREDEKYEEM 180
DB 121 dpqoqteccqkbcqrretprlmqtcqrcerryekekryeqoredeekyeem 180
OY 181 KEEDNKRDPQOREYEDCRRRCQOERROHCOLRCREORHGRGDMNPNRQSGSGRY 240
DB 181 keednkrdpqoreyedcrrrcceqrqecqrtrcrqesgprgqyqqrckelceeeeynqr 240
OY 241 EGEESQSDNPYFEDRSLSSTRFTEGHIISYLENFYGRSKLLRALKNRYLLEANPNA 300
DB 241 egeesqsdnpfyfderlsstrfteeeghlsvlenfygskllralknrylvleeanpna 300
OY 301 FYLPHTLADAILLVYGRGALKMIHNDRESYNLECGDVIRIPAGTTFYLLNRDNERL 360
DB 301 fylvphtladdaillvlygrgalkmihndresynlecgdviripagtlfyllnrddnerl 360
OY 361 HIAKFLQITSTPGQYKEFPAGGQNPETLSTFSKEILEALNTOTEKLRGVFGQOREG 420
DB 361 hiakflqitstpgqykefipaggnpepylstfskeilealntoteklrvgfvgqregv 420
OY 421 IIRASQEQIRELTRDSESRHWHIRGESSSRGPYNLFNKRPYLSNKYQAYEVKPEDYR 480
DB 421 iirasqeqireltrdseesrhwihirgesssrgpynlfnkrplysnkyqayevkpedyr 480
OY 481 QLODMLSLFIANYTQGSMMGPENTRSTKVVYVVASGEADVEMACHLSGRHGRGGGR 540
DB 481 qlodmlslvfiannytqgsmmgpntrstkvvyvvasgeadvemachlsgrhgrgggr 540
OY 541 HEEEDVHVEQVRLSRKEATVYLAGHPVVFSSGENULLFAGFINNONNENFLAR 600
DB 541 heeedvhveqvrlsrkeatvylaghpvfvssgenullfagfinnonnenflar 600
OY 601 ERNVLOQIEPQAMELAFAPRKEVESEFSNODSITFFPGPRHOQOOSPRSTKQOQPLVSI 660
DB 601 ernvloqiepqamela faaprkeveesfnodsqstiffpgprhqoqosprstktqoqplvsi 660
OY 661 LDFVGF 666
DB 661 ldfvgf 666

RESULT 2

ID W62829 standard; Protein; 666 AA.

AC W62829;

DT 27-OCT-1998 (first entry)

DE *Macadamia integrifolia* antimicrobial protein.
XX
XX antimicrobial protein; infestation; control.
XX
XX *Macadamia integrifolia*.
OS

Key Location/Qualifiers
FH peptide 1..28
FT /note= "signal peptide"
FT 29..666
FT Protein /note= "mature protein"

W09827805-A1.

02-JUL-1998.

22-DEC-1997; 97WO-AU00874.

20-DEC-1996; 96AU-0004275.

(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

WPI: 1998-377279/32.

N-PSDB; V42311.

PT Novel anti-microbial protein from e.g. *Macadamia integrifolia* -
PT useful for controlling microbial infestations of plants or mammals
XX
XX
PS Claim 1; Page 39-41; 96pp; English.
XX

CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
XX

Sequence 666 AA:

Query Match 96.3%; Score 3412; DB 19; Length 666;
Best Local Similarity 96.1%; Pred. No. 4.1e-288;
Matches 640; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

OY 1 MAINTSNLCSLLFLSLFLSTTVSLAESEFPDQREYEBCRQCMOLETSGQMRRCVSQCD 60
DB 1 maintsnlcsllflslflstvtslaeseftdqyeecrkqcmqltsgqmrtrcvsgcd 60
OY 61 KREEDIDMSKYDNOEDPQTECCQCRRCROESGPRQOQYCCORCKEICEEEYENRQR 120
DB 61 krtfeedidmskydngedpqtccqqrtrcrqesgprgqyqqrckelceeeeynqr 120
OY 121 DPQOQTECCQKHCQRRETPRHMQTCQRCERRYEKKRQOKRYEQOREDEKYEEM 180
DB 121 dpqoqteccqkbcqrretprlmqtcqrcerryekekryeqoredeekyeem 180
OY 181 KEEDNKRDPQOREYEDCRRRCQOERROHCOLRCREORHGRGDMNPNRQSGSGRY 240
DB 181 keednkrdpqoreyedcrrrcceqrqecqrtrcrqesgprgqyqqrckelceeeeynqr 240
OY 241 EGEESQSDNPYFEDRSLSSTRFTEGHIISYLENFYGRSKLLRALKNRYLLEANPNA 300
DB 241 egeesqsdnpfyfderlsstrfteeeghlsvlenfygskllralknrylvleeanpna 300
OY 301 FYLPHTLADAILLVYGRGALKMIHNDRESYNLECGDVIRIPAGTTFYLLNRDNERL 360
DB 301 fylvphtladdaillvlygrgalkmihndresynlecgdviripagtlfyllnrddnerl 360
OY 361 HIAKFLQITSTPGQYKEFPAGGQNPETLSTFSKEILEALNTOTEKLRGVFGQOREG 420
DB 361 hiakflqitstpgqykefipaggnpepylstfskeilealntoteklrvgfvgqregv 420
OY 421 IIRASQEQIRELTRDSESRHWHIRGESSSRGPYNLFNKRPYLSNKYQAYEVKPEDYR 480
DB 421 iirasqeqireltrdseesrhwihirgesssrgpynlfnkrplysnkyqayevkpedyr 480

Query Match	Best Local Similarity	Matches	604: Conservative	9: Mismatches	12: Indels	0: Gaps	0:
42	91.1%; Score 3227; DB 19; Length 625;	42	OCMOLEFGSGMARRCYSQCDKRFEEEDITMSKYDNGEDPOTECOCQCRRCRQESGFRPOQY	101			
1	96.6%; Pred. No. 4.5e-272;	1	gcmqlstsgqmrtrrcvsgcdkrfteedidwskydngedpqlccqgcqrctrcqsgsdbrqgy	60			
42	91.1%; Score 3227; DB 19; Length 625;	42	OCMOLEFGSGMARRCYSQCDKRFEEEDITMSKYDNGEDPOTECOCQCRRCRQESGFRPOQY	101			
1	96.6%; Pred. No. 4.5e-272;	1	gcmqlstsgqmrtrrcvsgcdkrfteedidwskydngedpqlccqgcqrctrcqsgsdbrqgy	60			

Qy	102	QGRCKEICEESEEENRNRDPQOQYEDQCQKCKRRREPRHMQTCOQRCERRLEKRRKQ	161
Qb	61	qrrrckelceeeeenrirdpbqqyveqcgkrckrrreleprlmhmqicqgrcerryekrkq	120
Qy	162	QKRYEEOOREDEDEKXEESGMKCEDNKNRDPQKQXEYEDCKRRCEQDEPRQHQCOLRCRPNOR	221
Db	121	qkryeeqqqedeekxyeemkcegdnnrdbpqgryedcrrhceqqeprlqygcqrrcqeqq	180
Qy	222	QHGRGDMWNPDRGSGSRRYEEGEEBQSDNPNYPFDEKSLSTRFFEECHIVLENFYGRSK	281
Db	181	qhgrgddlmnpdrgsgsryyeegeeqgsdnpylfderlsstrftrceghisvlenfygrsk	240
Qy	282	LRLAKNRLVLEENPNAPVLPRLHLDADAILVIGRGALKMIHMDNRESYLNCEGDVY	341
Db	241	lrlaknryrlvleeanpnatfvlprhldadailvlvgrrgalkmihmdnresynlcegdvi	300
Qy	342	RIPACSTFVYLIRDNRRNERLHAKFLQTSSTPGQYKKEFPFPGGONPEPYLSTFSKEILEAA	401
Db	301	ripagltfvlirndnmerlhakflqlstspgqykelfpaggqnepylstfskelleaa	360
Qy	402	LMFTQTEKLRGVFGQOREGVIIIRASQEOIRELJTRDSDSRRHWHLRGGESSRGPYNLFNKR	461
Db	361	lmtqterlrgvlygqregvliirasqeqrreltrdssertrwhlrggessrtpynlfnkr	420
Qy	462	PLYSKKYQGANVVRPEDRYRQJQMDLSVFIVNTVQSGMGPFFTRSTKVVVYVAASEGADV	521
Db	421	plyskkyqgayevvrpedryrqjdmcdvsfivantlqgsmgpftrfstkvvasseadv	480
Qy	522	EWACPHILSGRHGRCRGCGKRHEEEDVHYEQVRAFLSKREAIIVYLAGHPVYVSSGNELL	581
Db	481	ewacphilsgrrhgrrcgckrheeeedvhyeqvraliskreaiivylaahpvyvssgnell	540
Qy	582	LEAFGINAQNHNENFLAGERNVLQOJIEPQAMELAFAPRKEVEESFNQDOSIFPPGR	641
Db	541	lfafiginagqnhenflagrenvliqjlepgamelafaasrkeveelnsqdesiffppgr	600
Qy	642	QHQQSPSTSKQOQPLVSIIDENGF	666
Db	601	qhqqspstskqqpvlsvlaidvgr	625
RESULT 4			
R20181			
ID	R20181 standard; Protein; 566 AA.		
XX			
AC	R20181;		
XX			
DT	16-APR-1992 (first entry)		
XX			
DE	Sequence encoded by 67 kd T. cacao protein cDNA.		
XX			
KW	Cocoa; flavour; vicillin; seed storage protein.		
XX			
OS	Theobroma cacao.		
XX			
PN	WO9119801-A.		
XX			
PD	26-DEC-1991.		
XX			
PF	07-JUN-1991; 91MO-GB00914.		
XX			
PR	11-JUN-1990; 90GB-0013016.		
XX			
PA	(MNSC) MARS UK LTD.		
XX			
PI	Spencer ME, Hodge R, Deakin EA, Ashton S;		
XX			
DR	WPI, 1992-024418/03.		
XX			
DR	N-PSDB; Q20377.		
XX			
Recombinant cocoa proteins - are responsible for flavour in cocoa beans and produced in large quantities using yeast and bacterial			


```

Db      265 hggdkgllvrateeqretlrhassegghpwhlpfge-sigpyslldqrpshlanqhq 323
OY      471 AYEKPEDYRQLODMDLVFIAVITANVTOGSMWGPFTNSTKVVVVVVASGEADVEMACPRLSG 530
Db      324 lyeadaersfhdlehdsvsfanitagmsaplfntsfklayvpngkyaaelvcphrgs 383
OY      531 RHG----GRGGKRHEEEDVH-----YEOVARLRSKREAIIVLAGHPVFEVSSG 576
Db      384 qggeserdkgriseeeseeseegeaagqyhtlrlarispglafvypaghpftvaasr 443
OY      577 NEULLFARGINQNNHENTLAGRERNVLOQIERPAMELAFAPRKVEVSFSQDSIF 636
Db      444 dsnlqivclevhadrnekvltagad-nvllqldrvakalsfaskaevedevlgsrtekgf 502
OY      637 FPGPRO---HQOQSPRSTKQOQ 655
Db      503 lpgpeesggheregeeeere 524

RESULT      8
W62837
ID      W62837 standard; Protein: 637 AA.
XX
AC      W62837;
XX
DT      27-OCT-1998 (first entry)
XX
DE      Hordeum vulgare antimicrobial protein.
XX
KM      antimicrobial protein; infestation: control.
XX
OS      Hordeum vulgare.
XX
PN      WO9827805-A1.
XX
PD      02-JUL-1998.
XX
PE      22-DEC-1997; 97WO-AU00874.
XX
PR      20-DEC-1996; 96AU-0004275.
XX
PA      (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI      Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
DR      WPI; 1998-377279/32.
XX
PT      Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX
PS      useful for controlling microbial infestations of plants or mammals
XX
PS      Claim 1; Page 60-62; 96pp; English.
XX
CC      The sequence is that of an antimicrobial protein which can
XX
CC      be used to control microbial infestations in plants and mammalian
XX
CC      animals.
XX
SQ      Sequence 637 AA;

Query Match      24.4%; Score 864.5; DB 19; Length 637;
Best Local Similarity 35.1%; Pred. No. 1.1e-66;
Matches 222; Conservative 98; Mismatches 219; Indels 93; Gaps 20;

OY      70 SKYDNOEDPO--TECOQCORRCROESGPRQOQYCORCKEICEEVEEYNRQDPOQYE 127
Db      27 ashdeddrrghshsqcvtgrcgrer--pr---yshtarceq-----tdqgqh- 71
OY      128 QOQKHQCRRETERHMQTCQQRCEERYEKEKRQQRVEEQQRDEDEKYEERKKEEDNKR 187
Db      72 -----grhegeeggrgrwghgeereehgrgrwghgeereehgrgrwgh 119
OY      188 DPQGRREYDRCRCCEQEFERQOQHCOLRCREQORQHRG-----GDMANPQRGSGRXEBS 243

```

```

Db      120 gegereergrghrjgege-----reeergrgrghgegeereegrggrgeq 170
OY      244 E--EESQDN--PYVPERSLSTRFRTEEGHISVLENYGSKILRALKNRLVLEENPN 239
Db      171 erdeeqdsrrpvrffprstrrllqsdhfvralrpdqvsrlllgrlrvaimenpr 230
OY      300 AFVLPHTLADAILLYIGRGALKMTHDNRRESYNLECGDVIRIPAGTFFYLINRDNNR 359
Db      231 afvvpftdddygyvaygegvltlengekrsytlvkegdvltvapslmhlanldgrtk 290
OY      360 LHLAKFLQITSTGQYKEFFPAGQNDPEPLSTFSEILEALNTOTEKRGVFGQO--- 416
Db      291 lvalkllhltsvpqkf-qfl-----svkpllaelskrvlraafltsderlrlfngrrqg 344
OY      417 ---REGVITRASQEOJRELTRDSE---SRHWHIRRGCESSRGPRVNLFNKRPLYSKMYGQ 470
Db      345 ektrsvsivraseeqqlrelteaegqghrwpdlppfrgdsrltlleqrpshlanhgr 404
OY      471 AYEKPEDYRQLODMDLVFIAVITANVTOGSMWGPFTNSTKVVVVVVASGEADVEMACPRLSG 530
Db      405 lyeadaersfhdlangdvrvavaniitpasmcapylntqsfklavilegegevgivcphl-g 463
OY      531 RHG----GRGGKRHEEED-----VHREOVARLRSKREAIIVLA 566
Db      464 reseserehkgrrreeeeddqrrgrseseeseeeqgrvetyrvatrsigatfvvp 523
OY      567 GHPVVFVSS--GNENLLFARGINAONNHENFLAGRERNVLOQIERPAMELAFAPRKVEV 624
Db      524 ghpyveissagssnlqvcvfeinaernewlaagr-nvlygklgsapageltfgrparev 582
OY      625 EESFNSQDSI-FPGPROHQOQSPRSTKQOQ 655
Db      583 gevfragdqdegfvagp---eqgsregegeqe 611

RESULT      9
Y15244
ID      Y15244 standard; Protein: 626 AA.
XX
AC      Y15244;
XX
DT      09-NOV-1999 (first entry)
XX
DE      Peanut allergen, Ara h 1, amino acid sequence.
XX
KM      allergy; immune response; transgenic; allergen; epitope;
XX
KM      Immunoglobulin E; Ig E; binding site; peanut.
XX
OS      Arachis hypogaea.
XX
PN      WO938978-A1.
XX
PD      05-AUG-1999.
XX
PE      29-JAN-1999; 99WO-US02031.
XX
PR      27-AUG-1998; 98US-0141220.
PR      31-JAN-1998; 98US-0073283.
PR      13-FEB-1998; 98US-0074590.
PR      13-FEB-1998; 98US-0074624.
PR      13-FEB-1998; 98US-0074633.
XX
PA      (SOSI/) SOSIN H.
PA      (UYAR-) UNIV ARKANSAS.
PA      (UYNY ) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
XX
PI      Bannon GA, Burks AW, Sampson HA, Sosin H;
XX
DR      WPI; 1999-479189/40.
XX
DR      N-PSDB; 206382.
XX
PT      Modified allergen with reduced IgE binding, useful for treating e.g.

```


DE Soybean beta-conglycinin protein sequence.
XX
XX Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
KM allergic reaction; soybean; beta-conglycinin.
XX
OS Glycine max.
XX
XX WO945961-A1.
PN
PD 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US05494.
PE
XX 12-MAR-1998; 98US-0077763.
PR
XX 11-MAR-1999; 99US-0077763.
XX
XX (UYAR-) UNIV ARKANSAS.
PA
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS,
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
XX
XX WPI; 1999-551218/46.
DR
XX
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction -
XX
XX
XX
PS Disclosure; Fig 33A-B; 193pp; English.
XX
XX The invention provides a tertiary structure for the peanut allergen
CC Ara h 1. The Ara h 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to
CC ensure that the allergen is not introduced into genetically modified
CC food. The present sequence represents a soybean beta-conglycinin protein
CC sequence.
CC
SQ Sequence 605 AA;

Query Match 23.5%; Score 831.5; DB 20; Length 605;
Best Local Similarity 30.3%; Pred. No. 7.6e-64;
Matches 204; Conservative 145; Mismatches 223; Indels 101; Gaps 17;

QY 11 LLEFLSLFLSTVTYSLAEEFDRQ--EYEECKKQCMQLTSGQMRRCVSGCDKRFEDID 68
DB 8 LlllgIvflasvsvfgiaywekknpkhkclqscnserdsytrngacharc-----n 59
QY 69 WSKYDNOEDPQTECCQOCRCRQOESGPRQOQYQCRRCKEICEEEERYNRQD---PQQ 125
DB 60 llkveckeckkgelprprprpqhpereqpgq-----keededeqprlpfprpq 110
QY 126 YEOCQKHQORRETE-PRHMOTCOQRCERRYEKKRKQOQRYEEOQREDEKYEERKKEED 184
DB 111 prqeeehqegregeewpr-----keekrgkqgseededeeqge---rqfp 154
QY 185 NKRDPOQREYEDRCRRCEQOEPHQOQCARCEQOQROHGRGDMNMPORGSGRYEEGE 244
DB 155 fprpppqke-----erneeede-----eqgreseesed-----selr 188
QY 245 EEOSDNPYYFDESRSLSTRFTEGHISVLENFYGSRSLRALKNYRLVLEAPNNAFVLP 304
DB 189 rhknknpflfgsnrfellfknqygrlvqrfqgrspqlnltidyrllefnskpnclllp 248
QY 305 TMLDADATLIVTGGGALKMIMHNDNRESYMLEGCDVIRIPACTTFVLLINDNNEERLIAR 364
DB 249 nhadadyllivlmgltalivnddrdsyrlqsgdalrpsgltylvvnpdmnenlrlit 308
QY 365 FLQTIPTPGQYKFEFFPAGGONPEPYLSTFSKELLEALNTQTEKLRGVF-----GQ 415
DB 309 laipvnpkprfesflisstragqsylyqisrnllesaydtkfeelnkvlfisreegqgqe 368

QY 416 QR--EGVILRASQEOIRELTRDSESRHWHIRRGESSRGPRYNLFNKRPLYSNKQOAYE 473
DB 369 qrlgesviveisketqlalaskraks-----rktissedkprfnlrstdplysnklygffe 424
QY 474 VKPEDYRQLODMDSVFRIANVTQGSMMKGPFFNTRSTKVVVAVASGEADVEMACPHLSGRHG 533
DB 425 ltpenkpqlrldldflslvdmnegalllpfnfskaivllylnegdanlelv----- 475
QY 534 GRGSGK-----RHEEEDVHYEQVRLSKREALIVVLAGHPVYVSSGNENLLFARGIN 588
DB 476 ---glkeqgqgeqgeqlvrvkyraelseqdlfvipagyvvv--natsnlffalgin 530
QY 589 AONNHENFLAGREERNVLOOTIEPQAMELAFAPARKVESEFNSODOSIFPPQROHOQOSP 648
DB 531 aennqrnflagsqdnvlsqdlpsqvgelafpsagavvkkllknqrksyfvdagpkkkkee 590
QY 649 RSTRKQOQPLVLSIL 661
DB 591 kytk--gplssil 601

Search completed: March 1, 2001, 15:46:56
Job time: 221 sec

